

# How odgcrnwi becomes crowding: Stimulus-specific learning reduces crowding

Anke Huckauf

Bauhaus-University Weimar, Germany



Tatjana A. Nazir

ISC-CNRS, Lyon, France



Processes underlying crowding in visual letter recognition were examined by investigating effects of training. [Experiment 1](#) revealed that training reduces crowding mainly for trained strings. This was corroborated in [Experiment 2](#), where no training effects were obvious after 3 days of training when strings changed from trial to trial. [Experiment 3](#) specified that after a short amount of training, learning effects remained specific to trained strings and also to the trained retinal eccentricity and the interletter spacing used in training. Transfer to other than trained conditions was observed only after further training. [Experiment 4](#) showed that transfer occurred earlier when words were used as stimuli. These results thus demonstrate that part of crowding results from the absence of higher level representations of the stimulus. Such representations can be acquired through learning visual properties of the stimulus.

**Keywords:** crowding, visual letter and word recognition, perceptual learning

**Citation:** Huckauf, A., & Nazir, T. A. (2007). How odgcrnwi becomes crowding: Stimulus-specific learning reduces crowding. *Journal of Vision*, 7(2):18, 1–12, <http://journalofvision.org/7/2/18/>, doi:10.1167/7.2.18.

## Introduction

Recognizing a visual character is impaired by the presence of nearby characters in the visual field—a phenomenon known as crowding or lateral masking. Crowding is particularly pronounced in the visual periphery and can be experienced when fixating the cross in the middle of the next line while concentrating on the two R's.

XRP                      +                      R

The cross is to be depicted in the center of the line, and both R's should be equally separated from the flanked R to the left, and the isolated R to the right. In spite of equal distance from fixation, the flanked R is much harder to see than the isolated R. The influence of nearby characters on the perception of a target becomes stronger as the distance between target and flanker decreases and crowding typically amplifies at larger retinal eccentricities (e.g., Bouma, 1970; Huckauf, Heller, & Nazir, 1999). Because eccentricity and intercharacter spacing are low-level visual features, crowding has been assumed to result from interferences during sensory processing (e.g., Estes, 1972; Wolford, 1975).

Studies that have investigated limits of crowding all seem to suggest that this phenomenon is difficult to negotiate. For instance, crowding effects cannot be compensated for by unlimited viewing time (Townsend, Taylor, & Brown, 1971) or by varying visual contrast between targets and flankers (Huckauf, & Heller, 2002; Pelli, Palomares, & Majaj, 2004). That is, manipulations

of sensory input do not seem to alter the phenomenon. Crowding does not benefit from cueing the target position, which indicates that processing of flankers cannot be suppressed (Huckauf & Heller, 2002; Pöder, 2006). Crowding effects even increase when targets and flankers belong to different categories (e.g., letters vs. letter-like nonletters; Huckauf et al., 1999; Styles & Allport, 1986), suggesting that target selection cannot be improved either.

This stability of crowding effects is particularly surprising when we take into consideration that flanked letters are well recognized when we read words. In fact, under nonoptimal visual conditions letters are even better perceived when they are part of a word than when they are displayed in isolation (Reicher, 1969; Wheeler, 1970). This “word superiority effect” is not an all-or-none phenomenon though, as it shows a smooth transition from (familiar) words to (unfamiliar) nonwords: Letter recognition is the better the more the string approximates a word. This transition must be due to learning.

If learning can improve letter recognition in words, it is thus legitimate to assume that learning will also improve letter recognition in random combinations of letters (cf. nonwords). As a matter of fact, crowding effects have been reported to be larger in children than in adults (Atkinson, Anker, Evans, Hall, & Pimm-Smith, 1988), which again suggest that learning may alter these effects in adults. However, few studies have directly investigated this issue (for a review of learning effects in peripheral vision, see Westheimer, 2001).

To our knowledge, only Wolford, Marchak, and Hughes (1988) have tested how training influences crowding. These investigators compared training effects on crowding (referred to as lateral masking) to those on backward

masking. In their lateral masking condition, target letters were flanked either by a column of five H's on each side of the target or by one H above and below the target. Their results showed that unlike backward masking, lateral masking hardly diminished over 8 days of training.

However, as the same flanking letter (the letter H) was used throughout the entire experiment, it is unclear whether the experimental procedure used by Wolford et al. (1988) was suitable to capture effects of training. In fact, learning effects might have already reached asymptote after the first training session (cf. ceiling effect) and performance might not have improved afterward. Moreover, participants were not provided feedback although learning effects are typically observed following feedback (e.g., Dill, 2002; Fahle, 2002). Hence, although the work of Wolford et al. suggests that learning does not alter crowding, this possibility cannot be entirely ruled out. This study was aimed at clarifying this issue.

## Experiment 1

In Experiment 1, participants were trained to identify flanked target letters. During training, target letters were always presented in combination with the same two flankers—one on each side of the target letter (e.g., the letter A was always presented in the string XAZ, the target letter B in VBG, etc.). To determine whether potential effects of training were due to improvements in performing the task (unspecific learning) or to improvements in recognizing the strings (specific learning), we assessed learning effects after training for trained as well as for novel stimuli.

## Methods

Thirty students took part in the experiment for course credits.

Fifty three-letter strings from the study of Huckauf and Heller (2002; Experiment 4) were used in the experiment. Strings were divided into two sets: Set I contained strings with target letters A to N, and Set II contained strings with target letters M to Z. Letter strings were presented in upper case in Windows ARIAL with font size 10. Width and height of the characters were  $3 \times 4$  mm, corresponding to a visual angle of  $0.38^\circ \times 0.51^\circ$  at a viewing distance of 45 cm. Targets were displayed at  $4^\circ$  and  $7^\circ$  of eccentricity. Flankers were presented with a spacing of  $1^\circ$  on each side of the target (measured from letter center to letter center). The stimuli were black on a light grey background resulting in a Michelson contrast of 1.05. The monitor (14-in. CRT, Philips 4CM4270) had a refresh rate of 60 Hz and a resolution of  $600 \times 800$  pixels. The experiment was controlled by Experimental Run Time System (Beringer, 1993) version 3.0. Experimental Run

Time System ensures the synchronization of exposure time and refresh rate. Participants viewed the screen binocularly. To ensure fixation, we presented stimuli randomly in the right or the left visual field. A keyboard was situated between the monitor and participants.

Participants were randomly assigned to one of three groups. Ten participants received no training (control group), whereas the other 20 participants were trained. Half of the trained participants received feedback concerning the target letter in that a beep-tone signaled a correct response (target group). The other half received feedback concerning the flankers (flanker group) by a tone signaling that a response corresponds to one of the flankers. Feedback consisted of a beep-tone of 2000 Hz presented for 300 ms immediately following the response. For the 10 participants of the target group, the tone was presented when they correctly identified the central letter, for the flanker group, the tone was presented when the reported letter was one of the flanking letters. Participants were informed about the respective feedback information before starting the experiment. During the test session, no feedback was provided.

During training, the respective strings were presented twice at each of the two eccentricities ( $4^\circ$  and  $7^\circ$ ) in each of the two visual fields, yielding a total of  $n = 25$  (strings)  $\times$  2 (eccentricities)  $\times$  2 (visual fields)  $\times$  2 (repetitions) = 200 trials. In each group, five participants were trained with Set I and the other five were trained with Set II. After training, all participants performed the same test. During this test, all 50 strings were presented once at each eccentricity in each visual field in random order.

A trial in training and test sessions was composed of the following steps. A fixation cross in the center of the string was displayed for 1050 ms. Participants were instructed to fixate until the cross disappeared from the screen. One second following fixation onset, the three-letter string was presented for 50 ms. The task was to identify the central letter of the string by pressing the appropriate key on the board. The dependent measure was percentage of correct responses.

## Results

Comparisons between trained and untrained participants were performed for untrained stimuli only. Data were analyzed using analyses of variance (ANOVAs) with two within-subjects factors eccentricity ( $4^\circ$ ,  $7^\circ$ ) and visual field (left, right) and one between-subjects factor group (control, target, flanker). Only for the training groups, comparisons between trained and untrained stimuli were performed.

To statistically control for ceiling and floor effects, we carried out ANOVAs based on the absolute values as well as on arc-sine-transformed data. Note that throughout the manuscript, descriptive data refer to the percentages of correct responses, whereas the reported  $F$  and  $p$  values of the ANOVAs refer to arc-sine-transformed data.

Differences between analyses on percentage correct and arc-sine-transformed data, when they occur, are mentioned in the text.

Figure 1 plots the mean percentages of correct responses in the test session. Results are presented separately for the three groups of participants for the trained and untrained strings. Recognition of untrained strings differed marginally between the three groups (control, 29.1%; target, 35.5%; flanker, 37.1%),  $F(2, 27) = 2.93$ ,  $p = .07$ . The main factor group did not interact with any other variable. There was a main effect of eccentricity,  $F(1, 27) = 111.45$ ,  $p < .001$ , in that performance at  $4^\circ$  was better (44.7%) than at  $7^\circ$  (24.7%). The main effect of visual field was only marginally significant (left visual field, 30.3%; right visual field, 39.1%),  $F(1, 27) = 4.13$ ,  $p = .052$ . Eccentricity interacted with visual field,  $F(1, 27) = 11.76$ ,  $p < .01$ , in that the right visual field advantage was larger at  $4^\circ$  than at  $7^\circ$ .

A comparison of the two groups that received training revealed a main effect of eccentricity with better performance at  $4^\circ$  than at  $7^\circ$ ,  $F(1, 18) = 153.56$ ,  $p < .001$ . The interaction between visual field and eccentricity was marginally significant,  $F(1, 18) = 3.88$ ,  $p = .064$ . Target recognition for trained strings (48.3%) was significantly better than for untrained strings (36.3%),  $F(1, 18) = 12.78$ ,  $p < .01$ . Performance did not differ between groups,  $F < 1$ , and the factor group did not interact with any other variable. In short, the trained groups showed a (unreliable) 7% increase of performance for untrained strings (i.e.,

unspecific learning) and a stable 12% increase of performance for trained strings (i.e., specific learning).

## Discussion

Contrary to the findings of Wolford et al. (1988), the results of Experiment 1 thus demonstrate that recognition of a flanked target letter can be improved by training. Part of this benefit can be attributed to a general improvement in performing the task (i.e., unspecific learning). However, given that performance improved significantly more for trained than for untrained letter sequences, learning effects seem also to result from the acquisition of specific knowledge about the trained strings.

Both training groups showed a comparable amount of learning. Because feedback for the flanker group informed participants about flanker identity but never about target identity (which was the required response), participants could not simply enhance their performance by learning to associate the global visual pattern (of all three letters) with the correct response. Learning within this group must therefore be due to an improvement in distinguishing individual letters within the chain. Note that the present learning effects were already obvious after 4 repetitions of the same 25 strings, which corroborates our speculations about the potential failure of capturing learning effects in the study by Wolford et al. (1988).

As training in our first experiment was limited to approximately 25 min, the likelihood of observing a generalized improvement in task performance due to unspecific learning was relatively low. Experiment 2 was thus aimed at verifying whether more prolonged training would result in a stronger benefit for novel stimuli.

## Experiment 2

In Experiment 2, training was increased to 2 hr a day for three consecutive days. In contrast to the previous experiment, however, on each trial, flankers were randomly assigned to the target such as to prevent participants from getting familiar with the strings. Potential learning effects can thus be clearly attributed to a general improvement in performing the task. To further increase the probability of capturing unspecific learning effects, in addition to letters, unfamiliar symbols served as stimuli because learning effects might be more evident with less familiar stimuli.

## Method

Methods were the same as in Experiment 1 except for the following differences. In the letter condition, 11 letters

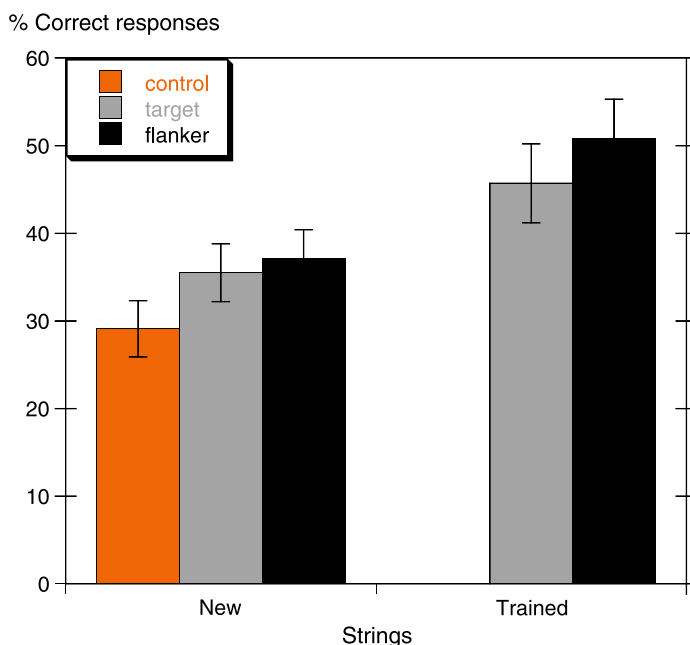


Figure 1. Mean percentage of correctly recognized target letters in novel and trained strings for the three groups receiving either no training (control), training with a beep-tone after correct target identification (target), and training with a beep-tone when one of the flankers was identified (flanker).



Table 1. Stimuli used in the symbol condition in Experiment 2.

of the alphabet were used as stimuli (C D F H J K L Q T V X). In the symbol condition, 11 symbols of the same width, height, and stroke width as the letters served as stimuli (see Table 1). Targets were displayed at 1°, 4°, and 7° of eccentricity. Each target was presented five times flanked by two different flanking letters chosen randomly at each trial from the remaining characters.

The experiment contained a total of 660 trials per participant per day: 2 character types (letters, symbols), 3 eccentricities (1°, 4°, 7°), 2 visual fields (left, right), 11 different targets, and 5 repetitions. Letters and symbols were presented in different experimental blocks. At each of the 3 days, participants started with the letter block. Within each block, trials were presented in random order. The dependent measure was percentage of correct responses. Participants were instructed to report the central character of the string by pressing predefined keys on the board. For this, 11 function keys were labeled with the respective letters or symbols. A break was introduced every 66 trials. The experimental session was preceded by an initial practice session consisting of 15 trials with letters not included in the experiment. Four male students took part in the experiment for course credits.

## Results

Figure 2 summarizes the results. On average, participants recognized 58.47% of the letters and 52.72% of the symbols. In the first block, overall performance for targets (letters and symbols) was 55.78%. In the second block it was 55.97%, and in the third block it was 55.03%. That is, there was no indication that performance improved over the three sessions, neither for letters nor for the symbols. Given the unambiguous data, inferential analyses were not performed.

## Discussion

The results of Experiment 2 show that when flankers changed randomly from trial to trial, performance for target characters does not improve with prolonged training. This observation is particularly interesting when we consider that in Experiment 1 performance for trained strings improved after only four repetitions of the same letter strings. The combined results of the two experiments thus suggest that recognition of flanked target characters can be improved, provided that targets and

flankers are learned together. Experiment 3 was designed to better discern conditions that promote these learning effects.

## Experiment 3

So far, our results seem to suggest that stimulus-specific learning underlies the observed improvements in recognizing flanked targets. On the basis of observations in visual word recognition, one could thus speculate that word-like memory traces are acquired through visual training with unfamiliar strings. These memory traces might be regarded as abstract entries in the orthographic lexicon, devoid of semantics and possibly also of phonology as proposed, for example, by Dehaene, Cohen, Sigman, and Vinckier (2005). According to these authors, part of the occipitotemporal “what” pathway is tuned to writing and forms a hierarchy of increasingly broader and more abstract local combination detectors. At the lower level, combinations of local oriented bars can form local shape fragment detectors with some tolerance over displacements and changes in size. At a subsequent stage, combinations of fragments can be used to form local shape detectors. This stage could thus detect letters, but only in a given case and shape. Abstract letter identities can be recognized at the next stage, by pooling activation from populations of shape detectors coding for

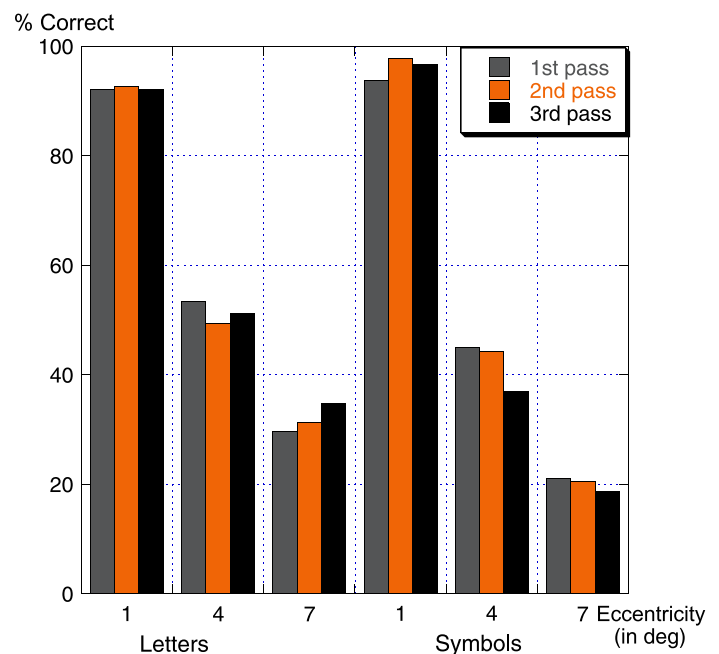


Figure 2. Mean percentage of correct responses for letters and symbols plotted as a function of eccentricity (1°, 4°, and 7°) and training session (first, second, and third day of training).



different visual versions of the same letter. If the acquisition of such abstract letter string representations explains the observed learning effects, low-level visual characteristics of the string should no longer affect performance once the string has been learned. In other words, although learning effects will be specific to the trained string, learning should nonetheless transfer when the string is displayed at eccentricities not used in training or when intercharacter spacing in strings is modified. To test this possibility, we trained strings in [Experiment 3](#) at a defined eccentricity and with a fixed intercharacter spacing in one of the two visual fields. During test sessions, the strings were probed either in the same conditions as during training, at a novel eccentricity, with another intercharacter spacing, or in the contralateral visual field. Note that if abstract letter string representations are indeed acquired during training, already existing representations might speed the learning process and facilitate transfer. Therefore, half of the strings were legal with respect to the German orthography (i.e., existing letter combinations in German), and half were illegal (i.e., letter combinations that hurt German orthography). The development of learning was probed at various stages during training because low-level visual characteristics of the strings may be acquired early during training, whereas abstract string characteristics may need additional time to develop.

## Method

Twenty-four students took part in the experiment for course credits.

Twelve letters were chosen as targets. Each target letter was embedded once in an orthographically legal string and once in an orthographically illegal string. All 24 strings were designed with the restrictions to not form meaningful words or common abbreviations.

The experimental method was identical to [Experiment 1](#) with the following exceptions: Only feedback concerning information about flankers was given because this condition guarantees that learning is not simply the result of associating a crowded display with the correct response. During training, all strings were presented at an eccentricity of 3° (half of the legal and illegal strings in one visual field and the other half in the opposite visual field). Each target letter was presented with the same number of trials in each visual field. Strings were presented in random order.

During test sessions, the same strings were presented at the trained position (identical), in the contralateral visual field (contralateral), at a smaller eccentricity of 1° (eccentricity), or with an enlarged spacing of 1° (inter-letter spacing is measured from center to center), yielding a total of 96 trials per participant in each test session. During test, no feedback was provided, and stimuli were displayed in random order.

To establish baseline performance, we run a first test session prior to training. The first training session comprised six presentations of each string (144 trials). After this short training, the test was performed a second time. In the subsequent long training session, each string was presented 30 times (720 trials). After training, a third test session was performed. In training and test sessions, participants' task was to identify the central letter of the string by pressing the appropriate key on the keyboard. Twelve of the 24 participants were retested 24 hr following training.

## Results

Absolute and arc-sine-transformed results in the test sessions were analyzed using a 3 (amount of training: baseline, after short training, after long training)  $\times$  2 (strings: legal, illegal)  $\times$  4 (test condition: identical, contralateral, eccentricity, spacing) MANOVA for repeated measures. There was a significant effect of training,  $F(2, 44) = 19.79$ ,  $p < .001$ . Whereas mean baseline performance was 58.70%, performance increased to 67.90% following short training, and to 75.10% following long training. Post hoc comparisons revealed that each condition differed significantly from the other two. There was a main effect of test condition,  $F(3, 66) = 65.64$ ,  $p < .001$ . Overall, mean performance for the identical condition (3° of eccentricity with normal spacing) was 64.80%. In the contralateral visual field performance dropped to 57.50%. At the smaller eccentricity it was 74.70%, and with enlarged spacing it was 71.90%. Except for the last two comparisons, all mean performances differed from each other. The data thus replicate the common finding that recognizing flanked letters is easier at smaller eccentricities and with larger spacing between letters. Training intensity and test condition interacted,  $F(6, 132) = 2.27$ ,  $p < .05$ . This interaction is depicted in [Figure 3A](#). To better visualize learning effects, [Figure 3B](#) plots increments over baseline after training with respect to performance in the baseline condition. As can be seen, the largest improvement occurred at the trained position, and the smallest improvement occurred when eccentricity or intercharacter spacing changed. Finally, orthographic regularity had no effect on performance (66.10% for illegal strings, 68.40% for legal strings),  $F(1, 22) = 1.91$ ,  $p = .18$ . None of the other factors were significant.

Performances for participants who performed the recognition test 24 hr after training are also depicted in the two figures. For these participants, an additional 4 (training intensity: baseline, after short training, after long training, 24 hr later)  $\times$  4 (test condition: identical, contralateral, spacing, eccentricity) MANOVA was performed. Like for the first analysis, test condition,  $F(3, 33) = 27.47$ ,  $p < .001$ , and training intensity,  $F(3, 33) = 5.53$ ,  $p < .01$ , produced significant main effects. Performance

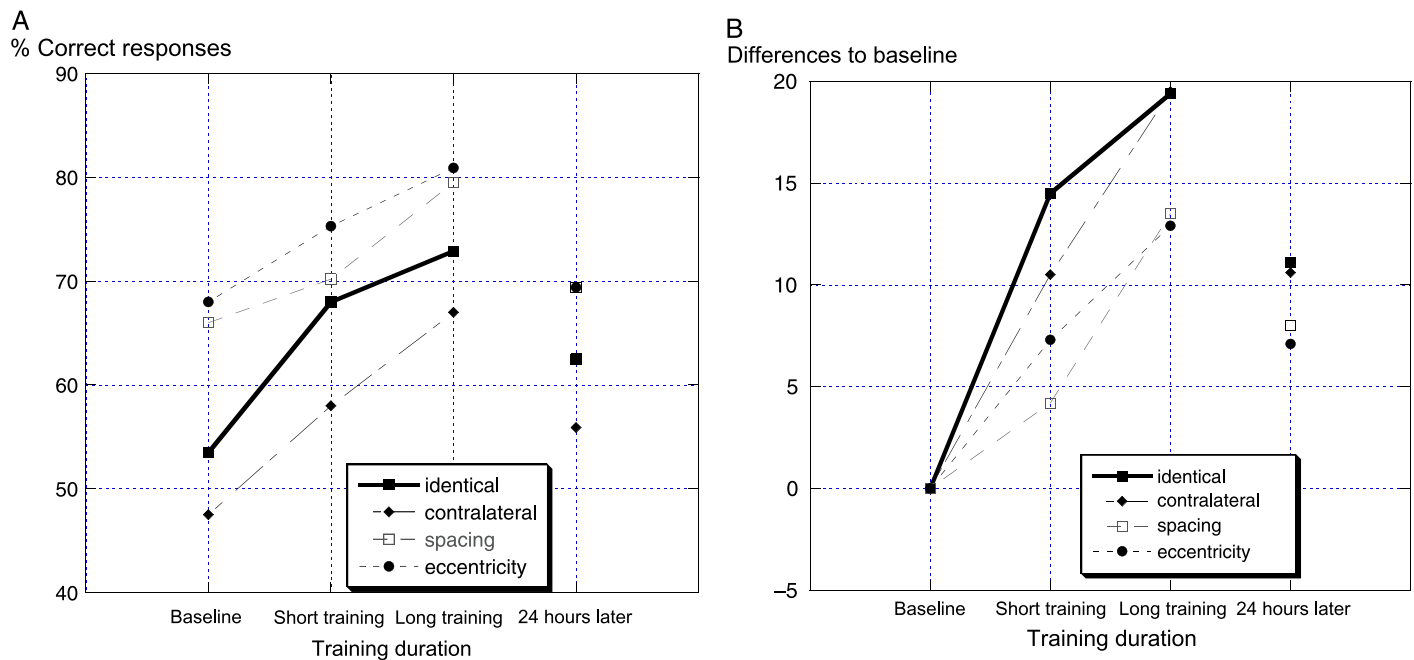


Figure 3. (A) Mean percentage of correct responses during four test sessions. The curves plot absolute values as a function of training duration for the trained manifestation of the strings (identical), for strings presented in the contralateral visual field, at a smaller than trained eccentricity, and with a larger than trained spacing. Note that the test session performed 24 hr after training involves only half of the total number of participants. (B) Increments over baseline are depicted.

in the fourth test session (64.30%) was significantly better than baseline (54.60%) but did not differ from performance after short (67.70%) or long training (75.30%).

## Discussion

Like the results of [Experiment 1](#), the present results demonstrate that recognition of targets in trained strings improves with training. Given that learning effects were largest when strings were presented at the trained eccentricity and with the trained intercharacter spacing, it is unlikely that learning evolved at abstract levels of stimulus representations only. This assertion is also corroborated by the fact that orthographic regularity did not produce significant effects. The results further show that performance in the contralateral field improved nearly as much as at the trained position. That is, retinal origins of the learning effect should also be ruled out. The slight difference between the two conditions might be due to differential cortical magnification of the three letters (Rovamo & Virsu, 1979).

Performance did not drop to baseline level even 24 hr after training. This indicates that learning did result in long-term memory traces, which are typically considered to be abstract in nature. However, despite the observed transfer of learning to untrained conditions, 24 hr following training the strongest learning effect was still

observed under conditions that exactly replicated training conditions. This stresses once more that learning was not restricted to abstract levels of stimulus representations.

The fact that orthographic legality did not affect performance was unexpected. On the one hand, this finding supports the notion that low-level visual characteristics of the strings are acquired during learning. On the other hand, however, effects of orthographic regularity are among the most robust effects in visual word recognition (e.g., Venesky & Massaro, 1979; Radach, Heller, & Inhoff, 2004). To corroborate the present findings, we therefore performed a further experiment in which learning effects in illegal nonwords were contrasted to those in words.

Until now, training effects were only studied in identification performance for flanked letters. To validate that the observed effects are indeed due to crowding, a comparison to isolated letters is still missing. This was added in [Experiment 4](#).

## Experiment 4

[Experiment 4](#) followed the procedure of the previous experiment. However, to clearly determine whether training reduces crowding effects, or whether it improves performance of trained stimuli in general, we additionally

presented isolated letters during training and test. If training reduces crowding, performance should mainly improve for flanked, but not for isolated letters. In addition, letters were also presented in a word context.

## Method

Eleven participants took part in Experiment 4. Ten target letters (A H K L M N O R S U) were trained either in isolation or embedded in nonwords and in common three-letter German words. Training and test conditions were the same as in Experiment 3.

## Results

Performances in the trained condition were analyzed using a 3 (amount of training: baseline, after short training, after long training)  $\times$  3 (context: isolated, word, nonword) MANOVA. This analysis detected a significant gain of performance from 60.60% at baseline to 69.80% and 75.40% after short and long training,  $F(2, 20) = 18.45$ ,  $p < .001$ . The effect of context was also significant,  $F(2, 20) = 113.01$ ,  $p < .001$ , in that performance for isolated letters (90.70%) was better than performance for letters in words (70.10%), which was better than performance for letters in nonwords (45.00%).

Moreover, context interacted with the amount of training,  $F(4, 40) = 5.77$ ,  $p < .01$  (Figure 4), showing that performance improved mainly for targets in words and nonwords, but not for isolated targets. To better understand

the data, for each context, we performed a separate ANOVA with the factors training (baseline, after short, after long training) and test condition (identical, contralateral, at a smaller eccentricity, with a larger spacing).

### Nonword context

In the nonword context, data replicated the effects observed in Experiment 3. Baseline performance was 28.2% for identical, 34.5% for contralateral, 39.1% for novel eccentricity, and 43.6% when intercharacter spacing increased. The amount of training,  $F(2, 20) = 15.11$ ,  $p < .001$ , and the test conditions,  $F(3, 30) = 8.24$ ,  $p < .001$ , produced significant main effects and a significant interaction,  $F(6, 60) = 2.62$ ,  $p < .05$ , which is depicted in Figure 5. Note that no gain of performance was obvious after the short training when the trained string was displayed at a novel eccentricity or with increased intercharacter spacing.

### Word context

In the word context, baseline performance was 50.9% for identical, 57.3% for contralateral, 70.9% for the novel eccentricity, and 60.9% when intercharacter spacing increased. Performances varied with the amount of training,  $F(2, 20) = 10.37$ ,  $p < .001$ , in that performance at baseline (60.00%) was lower than performance after the short (71.60%) and long training (79.10%). Also, the test condition produced significant differences in performance,  $F(3, 30) = 20.95$ ,  $p < .001$ . Contrary to nonwords, however, the two factors did not interact,  $F < 1$ , which

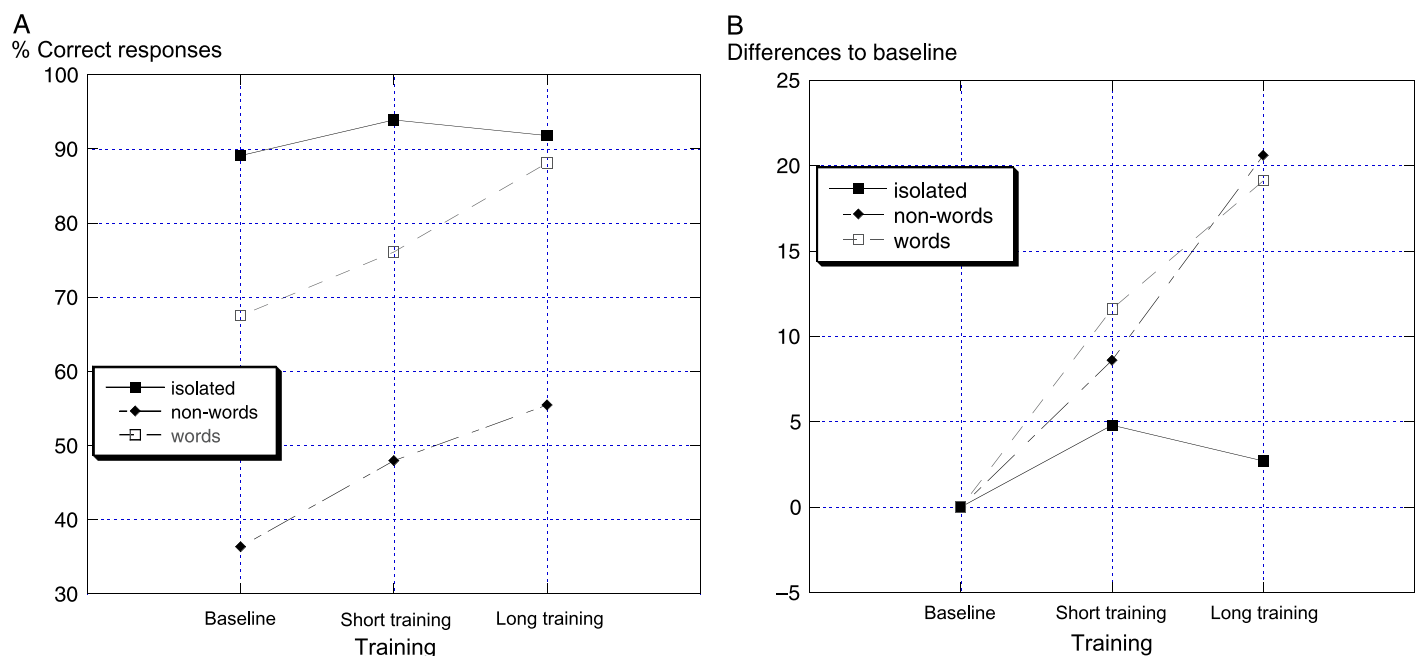


Figure 4. Mean performance (A) or increments over baseline (B) for words, nonwords, and isolated letters after short and long training in Experiment 4.

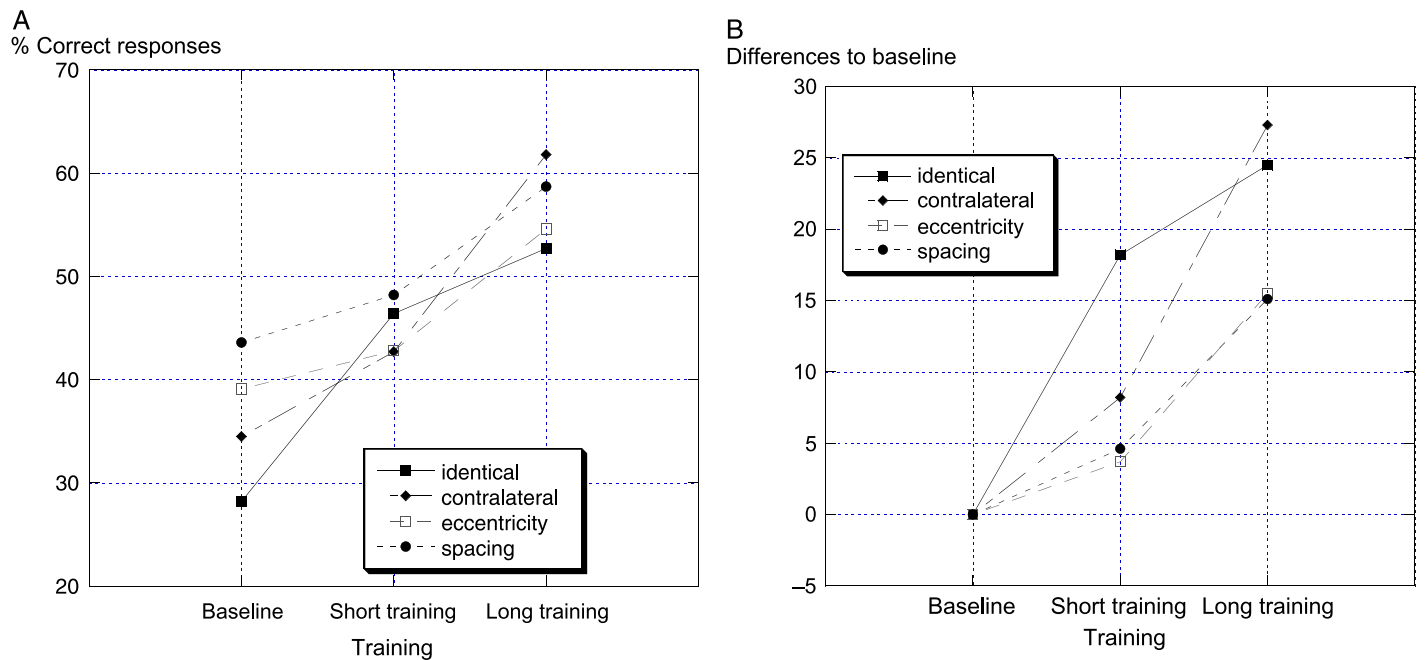


Figure 5. Mean performances (A) and increments over baseline (B) in nonwords sessions in Experiment 4. The curves denote the trained manifestation of the strings (identical), for strings presented in the contralateral visual field, at a smaller eccentricity, and with a larger spacing.

indicates that learning rates for the trained and untrained conditions did not differ (see Figure 6).

### Isolated letters

Performance for isolated letter was generally high but varied with training,  $F(2, 20) = 3.78$ ,  $p < .05$ . However, the

only significant difference to mean baseline performance (89.10%) was observed following short training (93.90%). After the long training, performance (91.80%) did not differ from either of the two other training conditions. The test condition also produced significant effects,  $F(2, 20) = 4.73$ ,  $p < .01$ . Hence, although performance at the trained (larger) eccentricity was 88.50%, at the smaller eccentricity

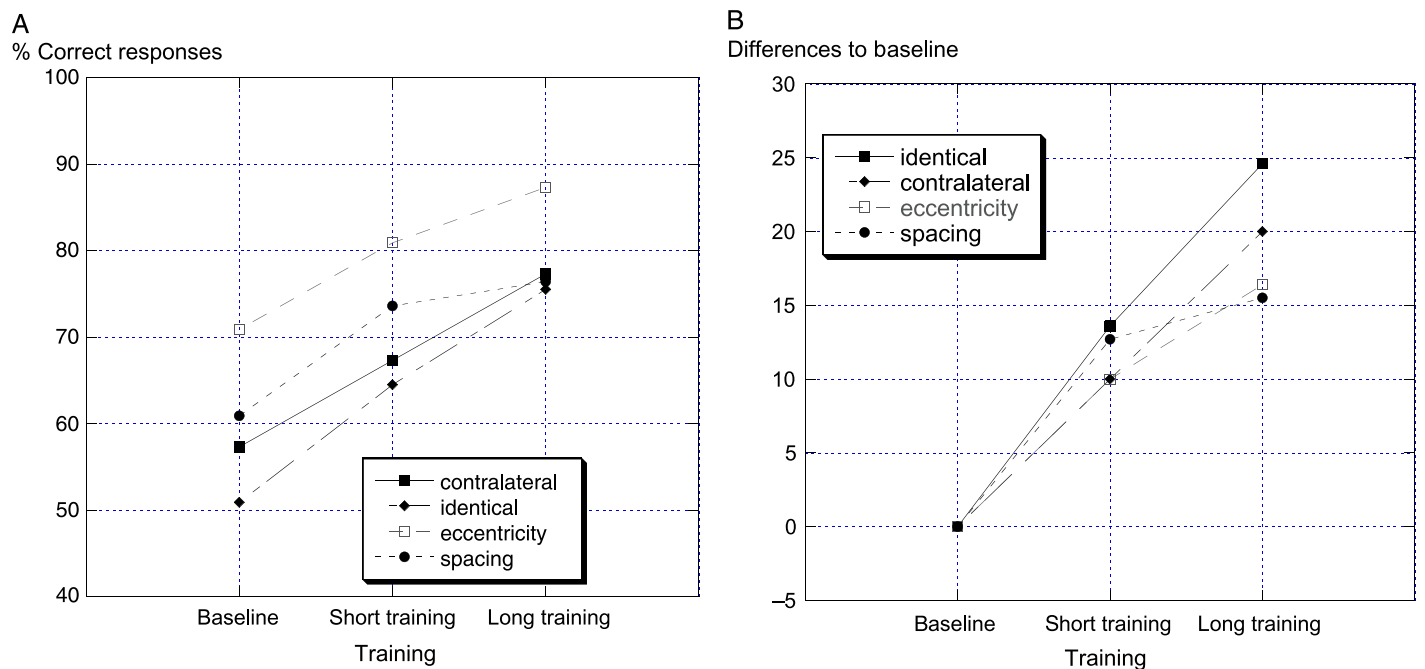


Figure 6. Mean performances (A) and increments over baseline (B) in words in Experiment 4. The curves denote the trained condition (identical), strings presented in the contralateral visual field, at a smaller eccentricity, and at a larger spacing.



it was 94.50%. In the contralateral visual field, performance was 91.20%. Training intensity and test condition did not interact,  $F < 1$  (see Figure 7).

### Effects on crowding

A summary of these data is given in Figure 8, which plots performance for the trained condition (identical) at baseline and after the long training as a function of context. Whereas at baseline, performance in the identical condition was 88.00% for isolated letters and 35.50% for letters in nonwords—that is, a crowding effect of 52.50%—after training, performance for isolated letters increased to 91.80% whereas performance in the nonword context increased to 55.50%—that is, a crowding effect of 36.30%. Hence, after 36 repetitions, crowding in trained letter strings was reduced by about one third. Interestingly, this small amount of training increased performance for letters in nonwords so much that it reached the level of performance for letters in words at baseline (58.40%).

### Discussion

The results of Experiment 4 suggest that learning reduces crowding: Whereas recognition of letters in strings was greatly improved by training, recognition of isolated letters barely varied and did not show a

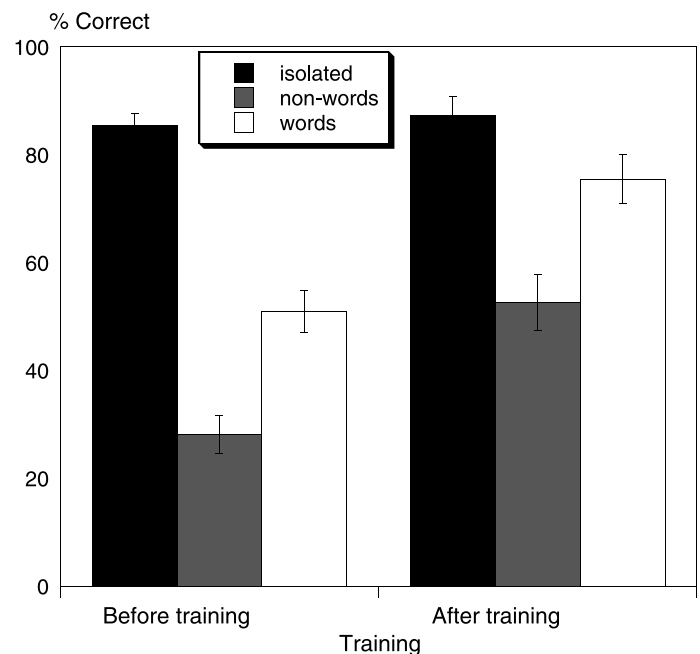


Figure 8. Mean percentage of correct responses for the trained manifestation (identical) of the strings measured in the test sessions at baseline (before training) and after the long training (after training). The differences between performance for isolated letters and letters presented in nonwords are referred to as crowding effect; the differences between targets presented in words and in nonwords are referred to as word superiority.

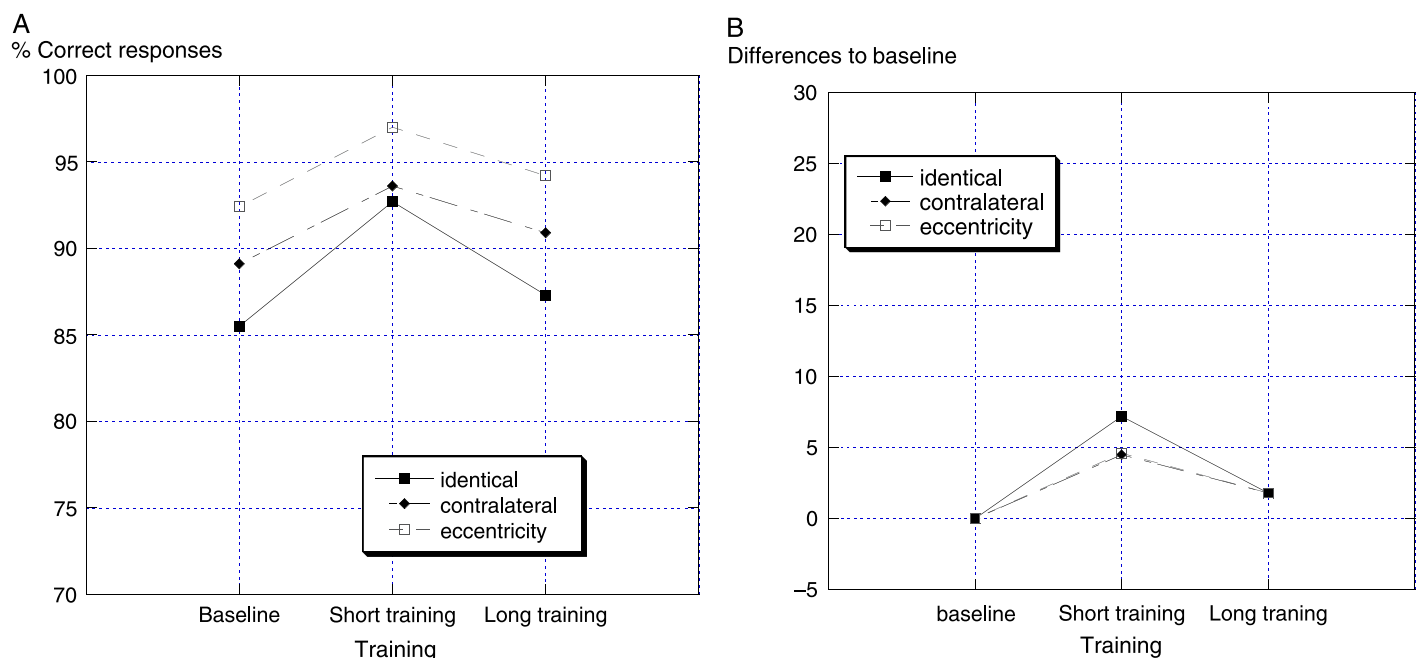


Figure 7. Mean performances (A) and increments over baseline (B) in isolated letters in Experiment 4. The curves denote the trained manifestation of the letters (identical), letters presented in the contralateral visual field, and at a smaller than trained eccentricity.

monotonic increase with training (see also Westheimer, 2001). This assertion has to be considered with caution though because ceiling effects cannot be entirely ruled out for isolated letters. Recall nevertheless that baseline performance for isolated letters was 88%, which leaves some space for potential training-related improvement (note also that ceiling effects have been statistically controlled for via arc sine transformation of the data). Moreover, given that training improved performance for highly familiar three-letter words as readily as for unfamiliar nonwords, the absence of learning for isolated letters cannot be ascribed to the high familiarity of the stimuli. Nonetheless, the assertion that learning was less efficient for isolated letters than for letters embedded in strings needs to be substantiated before more definite conclusions can be drawn (e.g., by equalizing baseline performance across conditions through the adjustment of stimulus parameters; see Farell & Pelli, 1999).

Results for nonwords replicated the effects observed in Experiment 3. That is, after short training learning effects were restricted to the trained eccentricity and to the trained intercharacter spacing. This superiority was also evident after the long training. However, after the long training, performance also improved for untrained conditions. For words, this transfer of learning to untrained display settings was already observed after the short training. In fact, performance for letters in words improved to the same amount in the trained and untrained display settings.

## General discussion

The present series of experiments was aimed at testing whether training can reduce crowding. Experiment 1 revealed that improvement in performance following training was largely restricted to the trained string (specific learning) although a small but unreliable improvement for untrained strings (unspecific learning) was also observed. When letter strings changed from trial to trial with no repetition (Experiment 2), no gain in performance was obvious after hours of training, even when less familiar stimuli than letters were used as targets. The absence of learning effects in Experiment 2 suggests that unspecific learning effects are of minor importance. Experiments 3 and 4 were then designed to identify conditions that promote stimulus-specific learning. Experiment 3 indicated that after short training, stimulus-specific learning was restricted to the trained retinal location and to the exact interletter spacing used in training. Following longer training, however, learning generalized to untrained eccentricities and spacing configurations, although this transfer was not complete. Performance with trained display settings remained superior to performance with untrained settings and this superiority was still obvious 24 hr following training. The

observed learning profile thus suggests that abstract representations of the stimulus (e.g., representations that are invariant to spacing and retinal eccentricity) and specific aspects of the letter strings were acquired through training. Finally, Experiment 4 showed that when words were used as embedding context for the target letter, performance improved equally for trained and untrained display settings. This latter experiment also showed that for letters presented in isolation, training effects were scarcely observable.

All together, the present results thus clearly show that training improves the ability to identify flanked letters and that this learning effect partly depends on the sensory characteristics of the strings during training (for similar observations regarding font, see Sanocki, 1987). Hence, unlike current accounts that promote only abstract visual representation of familiar orthographic strings (e.g., Deheane et al., 2005), in this study acquired knowledge about letter strings includes information about physical aspects of the stimulus—at least at the beginning of learning. With increasing experience stimulus representations seem to become more tolerant to surface variations, although it is yet unclear whether information about physical aspects of the stimulus ceases to be functional.

Whereas recognition of letters in strings improved after training, isolated letters did not profit from learning. Moreover, learning to recognize an embedded target letter improved even when participants received feedback about flanker identity only. To be able to report the correct target, training must thus have reduced interferences (pooling) between flankers and target. Taken together, the present findings suggest that learning reduces crowding. In fact, when confronted with a chain of letters, observers may actually attempt to process the string as a word, that is, holistically. For that purpose, visual information is spatially integrated. In the case of familiar chains, higher level internal representations of the stimulus exist already. Spatial integration of the chain might thus activate these representations and facilitate identification of the string and its embedded letters by top-down feedback. By contrast, in the absence of a higher level representation, holistic processing of the string results in interference among neighboring letter features, which is observed as the phenomenon of crowding (note that in tasks that do not encourage holistic processing like, e.g., visual search-like localization or detection tasks, interference between letters in unfamiliar strings differs from interference observable in identification tasks; Huckauf, 2006). The above speculation is in line with assumptions that conceptualize crowding as failure of feature integration (e.g., Bouma, 1970; Pelli et al., 2004; Wolford, 1975). However, it adds that this failure arises because of the attempt to process an unfamiliar chain of letters like a word (i.e., holistic processing) without top-down information because higher level stimulus representations are not available.

It has to be stressed although that there are several critical issues that constrain the current findings. First,

only skilled readers participated in this study. The limit of using this population is that feature- and letter-level information contained in the strings has already undergone extensive training. Hence, although the perception of unfamiliar visual objects might indeed benefit from training, skilled readers are not adequate for capturing such potential effects for the perception of isolated letters. A next important experimental step is therefore the investigation of training effects on the perception of isolated letters and on unfamiliar visual patterns in beginning readers.

Second, as already mentioned above, one problematic issue is surely the comparison of increments in correct responses across different levels of performance (e.g., learning to identify embedded versus isolated letters). Therefore, a replication of the present findings with an alternative measure of learning is desirable. One such solution would be to measure the threshold value of a given stimulus parameter because threshold of physical parameters provides a better estimation of the underlying metrics than proportion of errors (e.g., Farrell & Pelli, 1999). For such a measurement, however, one must take into consideration that each physical parameter is potentially subject to specific learning.

Third, to link the data to visual word recognition, effects of phonological and semantic information have to be taken into consideration. In fact, because in this study orthographic information was presented only in the visual periphery, the role of orthographic information and thus of sensory specificity could have been overestimated. The kind of perceptual learning that is reported here could also turn out to be of little relevance to natural reading because words are exposed at various retinal locations when a reader scans a page of text. This “training” might result in more position-independent internal representations and could explain why learning effects in the word context were more invariant to manipulations of spacing and retinal position. Curiously, however, due to the way eye movements are programmed during reading, most of the time words are perceived/fixated at the same retinal location (i.e., at the “preferred viewing position” slightly left of word center; e.g., Nazir, Ben-Bounayad, Decoppet, Deutsch, & Frost, 2004; Nazir, Heller, & Sussmann, 1992; Rayner, 1979; Vitu, O’Regan, & Mittau, 1990). Moreover, word recognition is effectively best at this preferred viewing position and drops with every letter of deviation from the “trained” retinal position (Nazir, 2000). This viewing position effect in word recognition is already observed after a few months of reading instructions (Aghababian & Nazir, 2000), which suggests that learning processes like those shown in this study may indeed underlie rapid, skilled word recognition. If this latter assumption is correct, the present data would thus have an important impact on current methods of reading instructions: Instead of only focusing on the training of letters and on the developments of higher level abstract word representations, perceptual training of whole words

might lead to faster and stronger improvements of visual word recognition.

Finally, it is worth noting that for object recognition in general, there is evidence that internal presentations of even familiar objects are not completely devoid of positional information. Objects are processed faster and more accurately when they are presented in a canonical view (e.g., Blanz, Tarr, & Bülthoff, 1999; Tarr, 1995). This canonical view correlates with the standard viewpoint for an object, which seems to mainly depend on observers’ experience with this particular view (Blanz et al., 1999). Assuming similar mechanisms for letter/word recognition, this study is the first to attempt investigating the genesis of such canonical views for letter strings. Learning means that internal representations of letter strings emerge based on sensory experiences. The more frequent the sensory experiences are, the more abstract the internal representation becomes; that is, the more tolerant it will be to surface variations. The results of this study suggest that one important effect of learning is that crowding in letter strings decreases with increasing learning. In the reverse, this means that insufficient top-down information might be regarded as one basic source underlying crowding effects.

## Acknowledgments

Commercial relationships: none.

Corresponding author: Anke Huckauf.

Email: anke.huckauf@medien.uni-weimar.de

Address: Bauhaus-Universität Weimar, Faculty of Media, Bauhausstr. 11, 99423 Weimar, Germany.

## References

- Aghababian, V., & Nazir, T. A. (2000). Developing normal reading skills: Aspects of the visual processes underlying word recognition. *Journal of Experimental Child Psychology*, 76, 123–150. [PubMed]
- Atkinson, J., Anker, S., Evans, C., Hall, R., & Pimm-Smith, E. (1988). Visual acuity testing of young children with the Cambridge Crowding Cards at 3 and 6 m. *Acta Ophthalmologica*, 66, 505–508. [PubMed]
- Beringer, J. (1993). *ERTS: Experimental Run Time System, Version 3.00*. Frankfurt: Computer Software.
- Blanz, V., Tarr, M. J., & Bülthoff, H. H. (1999). What object attributes determine canonical views? *Perception*, 28, 575–599. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226, 177–178. [PubMed]
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, 9, 335–341. [PubMed]

- Dill, M. (2002). Specificity versus Invariance of Perceptual Learning. In M. Fahle & T. Poggio (Eds.), *Perceptual learning* (pp. 219–231). Cambridge, MA: MIT Press.
- Estes, W. K. (1972). Interactions of signal and background variables in visual processing. *Perception & Psychophysics*, *12*, 278–286.
- Fahle, M. (2002). Learning to perceive features below the foveal photoreceptor spacing. In M. Fahle & T. Poggio (Eds.), *Perceptual learning* (pp. 197–218). Cambridge, MA: MIT Press.
- Farrell, B., & Pelli, D. G. (1999). Psychophysical methods, or how to measure a threshold and why. In R. Carpenter & J. G. Robson (Eds.), *Vision research: A practical guide to laboratory methods* (pp. 130–136). New Oxford University Press.
- Huckauf, A. (2006). Task set determines the amount of crowding. *Psychological Research*, paper still in press. [PubMed]
- Huckauf, A., & Heller, D. (2002). Spatial selection in peripheral letter recognition: In search of boundary conditions. *Acta Psychologica*, *11*, 101–123. [PubMed]
- Huckauf, A., Heller, D., & Nazir, T. A. (1999). Lateral masking: Limitations of the feature interaction account. *Perception & Psychophysics*, *61*, 177–189. [PubMed]
- Nazir, T. A. (2000). Traces of print along the visual pathway. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 3–23). Oxford, England: Elsevier.
- Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and recognition of printed words. *Brain and Language*, *88*, 294–311. [PubMed]
- Nazir, T. A., Heller, D. & Sussmann, C. (1992). Letter visibility and word recognition: The optimal viewing position in printed words. *Perception & Psychophysics*, *52*, 315–328. [PubMed]
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12):12, 1136–1169, <http://journalofvision.org/4/12/12/>, doi:10.1167/4.12.12. [PubMed] [Article]
- Pöder, E. (2006). Crowding, feature integration, and two kinds of “attention.” *Journal of Vision*, *6*(2):7, 163–169, <http://journalofvision.org/6/2/7/>, doi:10.1167/6.2.7. [PubMed] [Article]
- Radach, R., Heller, D., & Inhoff, A. W. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology*, *16*, 27–51.
- Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception*, *8*, 21–30. [PubMed]
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 275–280. [PubMed]
- Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*, 1–16. [PubMed]
- Sanocki, T. (1987). Visual knowledge underlying letter perception: Font-specific, schematic tuning. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 267–278. [PubMed]
- Styles, E. A., & Allport, D. A. (1986). Perceptual integration of identity, location and colour. *Psychological Research*, *48*, 189–200. [PubMed]
- Tarr, M. J. (1995). Rotating objects to recognize them: A case study of the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin and Review*, *2*, 55–82.
- Townsend, J. T., Taylor, S. G., & Brown, D. R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, *10*, 375–378.
- Venesky, R. L., & Massaro, D. W. (1979). The role of orthographic regularity on word recognition. In L. Resnick, & P. Weaver (Eds.), *Theory practice of early reading*. Hillsdale, NJ: Erlbaum.
- Vitu, F., O’Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics*, *47*, 583–600. [PubMed]
- Westheimer, G. (2001). Is peripheral visual acuity susceptible to perceptual learning in the adult? *Vision Research*, *41*, 47–52. [PubMed]
- Wheeler, D. D. (1970). Processes in Word Recognition. *Cognitive Psychology*, *1*, 59–85.
- Wolford, G. (1975). Perturbation model for letter identification. *Psychological Review*, *82*, 184–199. [PubMed]
- Wolford, G., Marchak, F., & Hughes, H. (1988). Practice effects in backward masking. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 101–112. [PubMed]